

Capacity of Text Marking Channel

Steven H. Low and Nicholas F. Maxemchuk

Abstract—We have proposed earlier watermarking text documents by slightly shifting certain text lines. Such a text line represents a noisy channel, and marking represents the transmission of a signal through this channel. The power of the signal represents the size of the shift and must be small for the marks to be imperceptible. In this paper, we formulate the channel capacity under a constraint on individual signal power. We show that to achieve the capacity, the shifts should be normally distributed, have maximum power, and adjacent shifts should be negatively correlated.

Index Terms—Channel capacity, copyright protection, text watermarking.

I. INTRODUCTION

MANY image watermarking methods have been proposed to protect motion and still pictures (see, e.g., [5], [6] and references therein). Most of these works, however, are suitable for images with rich greyscale. For text documents, whose images are often binary, marks placed by some of these methods can be easily removed by binarization. We have proposed in [1], [7], and [2] a different approach for marking text that is immune to the binarization attack. In our approach, a text line may be shifted slightly up or down, e.g., by one or two pixels, to encode a “1” or a “0.” This method turns out to be remarkably robust against severe distortions introduced by processes such as printing, photocopying, scanning, and lossy compression as in facsimile transmission [7]. Text watermarking using *any* technique, including ours, can always be defeated by retyping the document, possibly with the help of character recognition devices. For possible applications, attacks, and countermeasures, see [2]. Each line that is marked represents a noisy channel, and marking represents the transmission of a signal through this channel. In this paper, we consider the channel capacity when detection is made using text line centroids [7].

In Section II, we present a simple model of document profile, marking, and noise process. In Section III, we describe the centroid measurements that will be available for detection. In Section IV, we formulate our problem as determining the capacity of K parallel Gaussian channels with correlated noise under individual signal power constraints. The signal represents line shifts, and the constraint represents the requirement that the shifts be imperceptible. We show that to achieve the capacity, the shifts should be Gaussian distributed with *adjacent* shifts negatively correlated.

Manuscript received August 12, 2000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. R. L. de Queiroz.

S. H. Low is with the Computer Science and Electrical Engineering Departments, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: slow@caltech.edu).

N. F. Maxemchuk is with the AT&T Laboratories Research, Florham Park, NJ 07932 USA (e-mail: nfm@research.att.com).

Publisher Item Identifier S 1070-9908(00)10196-8.

II. PROFILE AND MARKING

Upon digitization, the image of a page is represented by a function $f(x, y) \in [0, 1]$, $x = 0, 1, \dots, W$, $y = 0, 1, \dots, L$, where W and L , whose values depend on the scanning resolution, are the width and length of the page, respectively. The value of $f(x, y)$ represents the greyscale at position (x, y) . The horizontal profile $h(y) = \sum_{x=0}^W f(x, y)$ is the projection of $f(x, y)$ along the horizontal direction. It measures the total intensity along the horizontal scan line y .

Lines that are marked (shifted) are always surrounded by two control lines that are not shifted. Marking shifts the middle profile block slightly to the left (up) or right (down) while leaving the neighboring control blocks unchanged. We *assume* that, after compensating for the structural distortions estimated from the control blocks, the remaining distortion can be modeled by additive white Gaussian noise (see [7] for an empirical justification). Hence, if the profile of an uncorrupted document is $h(y)$, then that of the corrupted document is $h(y) + G(y)$, where $G(y)$ are zero-mean independent Gaussian random variables with variance σ^2 . If the image is binary, i.e., $f(x, y) \in \{0, 1\}$, then the Gaussian noise assumption amounts to that the profile $h(y)$ and the standard deviation of the noise are large compared with one pixel, which seems reasonable.

III. CENTROID DETECTION

For the i th block defined in the region $[b_i, c_i]$ with profile $h(y)$, $y = b_i, b_i + 1, \dots, c_i$, its *centroid* is

$$c_i = \frac{\sum_{y=b_i}^{c_i} y h(y)}{\sum_{y=b_i}^{c_i} h(y)}.$$

When the profile is corrupted by additive white Gaussian noise $G(y)$ to become $h(y) + G(y)$, the centroid becomes

$$c'_i = \frac{\sum_{y=b_i}^{c_i} y (h(y) + G(y))}{\sum_{y=b_i}^{c_i} (h(y) + G(y))}.$$

We define the difference $N_i := c'_i - c_i$ as the “centroid noise.” N_i are independent since $G(y)$ is white.

In Fig. 1, the even blocks c_{2i} carry information and are shifted by X_i , and the odd blocks are control blocks that are not shifted. X_i is a negative if block $2i$ is shifted to the left and positive otherwise. We perform differential detection by measuring the

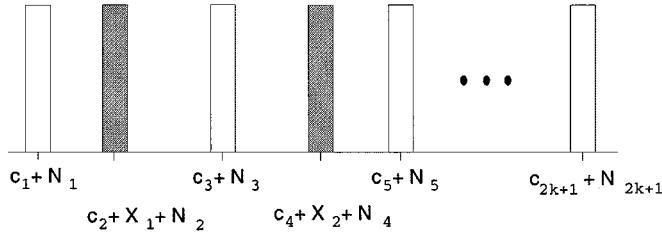


Fig. 1. K parallel marking channels: profile of a document page with K marked blocks (shaded). Here c_i is the centroid of block i on the original unmarked and uncorrupted copy, X_i is the shift of block i on a marked copy, and N_i is the centroid noise on a corrupted copy.

distances between the marked block and the neighboring control blocks

$$L_i := (c'_{2i} - c'_{2i-1}) - (c_{2i} - c_{2i-1}) = X_i + (N_{2i} - N_{2i-1}) \quad (1)$$

$$R_i := (c'_{2i+1} - c'_{2i}) - (c_{2i+1} - c_{2i}) = -X_i + (N_{2i+1} - N_{2i}) \quad (2)$$

and using the difference $Y = (Y_i = L_i - R_i, i = 1, \dots, K)$ to make a decision. Differential decoding eliminates the effects of translation, a common distortion in photocopying and facsimile transmission.

IV. MARKING CHANNEL CAPACITY

In [7], we have derived the distribution of the centroid noise N_i and have shown that, for a typical profile, it is accurately approximated by a Gaussian distribution with zero mean and with variance

$$\nu_i^2 = \frac{\sigma^2 w_i}{H_i^2} (\delta_i^2 + (w_i^2 - 1)/12) \quad (3)$$

where

$H_i = \sum_{b_i}^{e_i} h(y)$ is the “weight” of block i ;
 $w_i = e_i - b_i + 1$ is the width of block i ;
 $\delta_i =$ is the deviation of the centroid from the
 $c_i - (e_i + b_i)/2$ middle of block i .

Then from (1) and (2), the decision variable Y is given by $Y_i = S_i + Z_i$, $i = 1, \dots, K$ where the signal $S_i = 2X_i$. Here, the noise $Z_i = 2N_{2i} - N_{2i-1} - N_{2i+1}$ are zero-mean Gaussian. Hence, we have K parallel Gaussian channels whose inputs are vector $S = (S_1, \dots, S_K)$, and whose outputs are vector $Y = (Y_1, \dots, Y_K)$. The noise vector $Z = (Z_i, i = 1, \dots, K)$ has a subdiagonal covariance matrix C_Z

$$(C_Z)_{ij} = \begin{cases} 4\nu_{2i}^2 + \nu_{2i-1}^2 + \nu_{2i+1}^2, & \text{if } i = j \\ \nu_{2i+1}^2, & \text{if } j = i + 1 \\ \nu_{2j+1}^2, & \text{if } i = j + 1 \\ 0, & \text{else} \end{cases} \quad (4)$$

where ν_i^2 is given by (3).

Our goal is to choose the signal S so as to maximize the mutual information $I(S; Y)$ of the K parallel channels subject to individual power constraint $\max_{F_S} I(S; Y)$ subject to $ES_i^2 \leq p_i$, $i = 1, \dots, K$. Here, the maximization is over

the joint distribution F_S of signal vector S . The maximum $\max_{F_S} I(S; Y)$ is called the channel capacity. This is similar to a classical problem where the aggregate signal power over all K channels is constrained (see, e.g., [4] and [3]).

From [3, Ch. 10.5], the capacity can be achieved by choosing the inputs S_i to be joint zero-mean Gaussian random variables with covariance matrix C_S , which maximizes the determinant of the matrix $C_S + C_Z$ subject to the power constraint. Hence, the problem is equivalent to finding a C_S that

$$\max_{C_S} \det(C_S + C_Z) \quad (5)$$

$$\text{subject to } (C_S)_{ii} \leq p_i, \quad i = 1, \dots, K \quad (6)$$

where $\det M$ denotes the determinant of a matrix M . Let the maximum value be A . Then the channel capacity is $(1/2) \log_2(A / \det C_Z)$ bits.

By Hadamard's inequality, $\det(C_S + C_Z) \leq \prod_i ((C_S)_{ii} + (C_Z)_{ii})$, with equality if and only if $C_S + C_Z$ is diagonal. Hence, the unique maximizer C_S^* of (5) and (6) has diagonal terms maximized and off-diagonal terms being the negative of those of C_Z .

From (6) and (4) we have

$$(C_S^*)_{ij} = \begin{cases} p_i, & \text{if } i = j \\ -\nu_{2i+1}^2, & \text{if } j = i + 1 \\ -\nu_{2j+1}^2, & \text{if } i = j + 1 \\ 0, & \text{else.} \end{cases}$$

Hence, to achieve the channel capacity, the line shifts should be Gaussian distributed with maximum power and adjacent shifts should be negatively correlated. Moreover, the correlation should be of the same magnitude as the noise variance of the centroid that separates the two shifts.

The maximum value of (5) and (6) is then $A = \prod_i (p_i + 4\nu_{2i}^2 + \nu_{2i-1}^2 + \nu_{2i+1}^2)$, where ν_i^2 given by (3) depends on the profile parameters. Hence, the channel capacity is

$$\frac{1}{2} \log_2 \frac{\prod_i (p_i + 4\nu_{2i}^2 + \nu_{2i-1}^2 + \nu_{2i+1}^2)}{\det C_Z} \quad \text{bits} \quad (7)$$

where C_Z is the subdiagonal covariance matrix given by (4).

We close by presenting an example from [7], where two pages of document were watermarked. The first page contained eight marked lines, and the second page contained eleven, for a total of $K = 19$ marked lines. The size of the shift used in the experiment was two pixels, so we take the power constraint to be $p = 4$ pixels² for all lines $i = 1, \dots, 19$. The centroid noise variance was empirically measured to be $\nu^2 = 0.0781$ pixels². Assume that all lines have the same centroid noise variance. Then the 19×19 noise covariance matrix is

$$C_Z = \begin{bmatrix} 6\nu^2 & \nu^2 & 0 & \dots & 0 \\ \nu^2 & 6\nu^2 & \nu^2 & \dots & 0 \\ & & & \dots & \\ 0 & 0 & 0 & \dots & 6\nu^2 \end{bmatrix}.$$

Using (7), the channel capacity is 31 bits.

REFERENCES

- [1] J. Brassil, S. Low, N. Maxemchuk, and L. O’Gorman, “Electronic marking and identification techniques to discourage document copying,” *IEEE J. Select. Areas Commun.*, vol. 13, Oct. 1995.
- [2] J. T. Brassil, S. H. Low, and N. F. Maxemchuk, “Copyright protection for the electronic distribution of text documents,” *Proc. IEEE*, vol. 87, pp. 1181–1196, July 1999.
- [3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [4] R. G. Gallager, *Information Theory and Reliable Communications*. New York: Wiley, 1968.
- [5] *IEEE J. Selected Areas Commun.: Special issue on Copyright and Privacy Protection*, vol. 16, May 1998.
- [6] *Proc. IEEE: Special issue on Identification and Protection of Multimedia Information*, vol. 87, July 1999.
- [7] S. H. Low, N. F. Maxemchuk, and A. M. Lapone, “Document Identification for Copyright Protection using Centroid Detection,” *IEEE Trans. Commun.*, vol. 46, pp. 372–383, Mar. 1998.